Health of offspring of subfertile couples

Kuiper, Derk

DOI:
10.33612/diss.92269971

Publication date:
2019

Neurodevelopmental and cardiometabolic outcome in 4-year-old twins and singletons born after IVF

D. Kuiper
A. Bennema
S. la Bastide-van Gemert
J. Seggers
P. Schendelaar
M.L. Haadsma
A. Hoek
M.J. Heineman
M. Hadders-Algra

Reproductive BioMedicine Online 2017; 34: 659-67

2

Abstract

This prospective cohort study evaluated whether the cognitive development, neurological condition, anthropometrics and blood pressure of 4-year-old IVF twins differed from those of 4-year-old IVF singletons; 103 IVF singletons and 48 IVF twins born after conventional IVF treatment were included. Primary outcome was total intelligence quotient (IQ). Secondary outcomes were minor neurological dysfunction, anthropometrics and blood pressure. Unadjusted analyses found that the total IQ score of twins was lower than that of singletons, with a mean difference of −5.4 (–9.7 to −1.0). Weight (singletons: 18.6 [18.1 to 19.1] kg; twins: 16.9 [16.0 to 17.9] kg) and height (singletons: 108.8 [107.9 to 109.8] cm; twins: 105.9 [104.0 to 107.7] cm) of twins were lower than those of singletons (mean values [95% CI]). All differences disappeared after adjusting for mediators and confounders. Neurological outcome, systolic and diastolic blood pressure of twins and singletons were similar. Four-year-old IVF twins had a lower total IQ (−5.4 points), lower bodyweight (−1.7 kg) and were shorter (−2.9 cm) than 4-year-old IVF singletons. After adjustment, the adverse twin effect disappeared, implying that increased risk for impaired health and development in twins also holds true for IVF twins, and is not altered by IVF.

Introduction

Worldwide, the number of twin pregnancies is gradually increasing, partly because of the steadily rising application of assisted reproduction techniques, such as IVF and intracytoplasmic sperm injection (ICSI).[1] In Europe, however, the general multiple birth rate has declined since 2000, because of increased use of elective single embryo transfer in IVF.[2] Nevertheless, in 2012, 17.9% (17.3% twins and 0.6% triplets) of all European IVF pregnancies resulted in multiple births, with large cross-county variation ranging from 5.2% in Iceland, 7.4% in the Netherlands, up to 37.2% in Lithuania.

The high number of multiple births has caused concern. In general, twin pregnancies are associated with worse perinatal outcome than singleton pregnancies.[3] In addition, perinatal outcome of IVF twin pregnancies might be worse than that of naturally conceived twin pregnancies; some studies have suggested that IVF twin pregnancies are associated with an increased risk of obstetrical complications, such as preterm birth, low birthweight, Caesarean section and perinatal mortality, whereas others did not.[4-8]

Despite the high number of multiple births in IVF pregnancies, and the increased risk of obstetric complications in IVF twins, only a few studies have addressed differences in developmental outcome after the perinatal period between IVF twins and IVF singletons. Bonduelle et al. reported that cognitive development at 2 years, as measured by the Bayley Scales of Infant Development of twins born after IVF or ICSI, was significantly worse than that of singletons born after IVF or ICSI.[9] It was not clear, however, whether this difference could be attributed to a higher prevalence of perinatal adversities in twins. Pinborg et al., who used parental questionnaires to study developmental outcome, found that 3–4-year old IVF/ICSI twins were more likely to receive speech therapy than IVF/ICSI singletons. This difference persisted when effect modification by birth weight was taken into account.[10] Pinborg et al. also reported that the prevalence of cerebral palsy and psychomotor retardation in IVF twins aged between 2 and 7 years was similar to that of IVF singletons.[11] To the best of our knowledge, no study has addressed differences in cardiometabolic health between IVF twins and IVF singletons in childhood. Yet, it is known that naturally conceived twins are lighter and shorter than naturally conceived singletons at 4 years and at 9 years.[12-14] Whether naturally conceived twins are at higher risk for cardiovascular morbidity than naturally conceived singletons is not clear. The study by De Geus et al. suggests that this is not the case: adult blood pressure of naturally conceived twins was similar to that of their naturally conceived singleton siblings.[15]

The above overview indicates that our knowledge on differences in developmental outcome and cardiometabolic health between IVF twins and IVF singletons is limited. Therefore, the aim of the present study was to evaluate whether cognitive development in intelligence quotient (IQ) scores, neurological outcome of minor neurological dysfunction (MND), anthropometrics, including skin fold thicknesses, and blood pressure of 4-year-old IVF twins differs from that of 4-yearold IVF singletons. We hypothesized that developmental outcome and cardiometabolic condition of IVF twins is worse than that of IVF singletons, but that this difference will disappear after adjustment for effect mediators such as gestational age at birth and being small for gestational age.[16-19]

Materials and methods

Participants

Participants were singleton and twin members of the Groningen ART cohort study and the multicenter (Amsterdam Medical Center and University Medical Center Groningen) PGS Follow-up Study. The Groningen ART Cohort Study is a longitudinal assessor-blinded follow-up study aiming to investigate the potentially independent effects of ovarian stimulation and the in-vitro laboratory procedures on the offspring's health and development. To this end, subfertile couples (couples who were not able to conceive within 12 months of the start of unprotected sexual intercourse) and couples who were not infertile were recruited at the Department of Reproductive Medicine of the University Medical Center Groningen between March 2005 and December 2006. The Groningen ART cohort study consists of three groups of children: children born after ovarian stimulation and IVF/ICSI, i.e., conventional IVF/ICSI; modified natural cycle IVF; and natural conception (for details see Middelburg et al., 2009).[20] In the present study only singletons and twins conceived after ovarian stimulation and IVF/ICSI were included. Note that, in previous publications, we only reported on health and development of the singleton infants of the Groningen ART cohort. The PGS Follow-up study is a randomized, double-blind, controlled trial in which outcome of children born after ovarian stimulation and IVF with preimplantation genetic screening (PGS) is compared with that of children born after ovarian stimulation and IVF/ICSI without PGS. Women meeting the inclusion criteria, e.g., age ranging between 35 and 41 years, and having had no previously failed IVF-cycles, were recruited for the PGS trial at the Department of Reproductive Medicine of the Amsterdam Medical Center or University Medical Center Groningen between May 2003 and November 2005 (for details see Mastenbroek et al., 2007).[21] In the present study, only children who were conceived without PGS were included. Because of the slightly different inclusion criteria of the two studies background (social, fertility, and neonatal) characteristics of the groups showed some differences. Compared with the Groningen ART study, in the PGS study, maternal and paternal age at conception was higher (maternal age; mean values, [SD]): PGS 37.7 [1.63]; assisted reproduction techniques 32.9 [3.23]; P = 0.001), maternal educational level was higher (PGS 56%; assisted reproduction techniques 38%; P = 0.032), children were more often conceived with the help of ICSI (PGS 44%; assisted reproduction techniques 64%; P = 0.019) and more often had a low Apgar score (<7) (PGS 54%; assisted reproduction techniques 0%; P < 0.001). Other background characteristics did not differ between the two studies (data not shown).

In both studies, children born after oocyte or embryo donation and after oocyte cryopreservation were excluded. All participating children were born to subfertile parents, i.e., couples who were not able to conceive after 12 months of unprotected sexual intercourse. Four-year-old outcomes of the Groningen assisted reproduction technique cohort singletons born after ovarian stimulation and IVF and that of the PGS singletons conceived without PGS have been reported previously.[22-25]

A total of 89 neonates born as a result of ovarian stimulation and IVF/ICSI were recruited in the Groningen ART cohort study (26 twin infants and 63 singletons); 62 ovarian stimulation and IVF/ICSI neonates without PGS in the PGS study (22 twin infants and 40 singletons). Attrition at the age of 4 years was 11% in the Groningen ART cohort study (7% in singletons, 19% in twins), 17% in the PGS study (15% in singletons, 21% in twins), resulting in an overall attrition of less than 15%. As a result, 103 singletons and 48 twins born after ovarian stimulation and IVF/ICSI participated in the follow-up at 4 years. Attrition among singletons was non-selective in social and perinatal background factors. Attrition in twins, however, was selective: couples of twin pregnancies who dropped out of follow-up had a lower education than couples of twins retained in follow-up (prevalence of high education: maternal, drops outs 0%, retained 54% (P = 0.001); paternal, drop outs 17%, retained 53% (P = 0.048). In addition, twin dropout was associated with maternal smoking during pregnancy and low birthweight (maternal smoking: attrition 17%, included in follow-up 0%; P = 0.040; birth weight: attrition 2074 g, included in follow-up 2480 g; P = 0.009). Of the 48 twins who participated in the study 44 (92%) were born after double embryo transfer. Information on the number of embryos transferred in the remaining two twin pregnancies was lacking.

Information on socioeconomic status and the prenatal, perinatal and neonatal period was collected on standardized charts. High level of education was defined as a higher vocational education or university education. Information on time to pregnancy (TTP), a proxy for the severity of subfertility, was retrieved from the medical files. Note, that in case of miscarriage TTP can be less than 1 year, as TTP has a new onset. Parents provided written informed consent. The Medical Ethics Committee of the UMCG approved the study design of the follow-up at 4 years of both the Groningen ART cohort study and the PGS Follow-up study on 7 July, 2009 (reference number M09.074824).

Follow-up assessment

The follow-up examination at the age of 4 years took place at the University Medical Center Groningen or, when a family was not able to come to the University Medical Center Groningen for logistic reasons, at the child's home. The examination of the children was carried out by trained researchers who were blinded to the mode of conception of the children. It consisted of the assessment of neurological and cognitive development and the evaluation of anthropometrics and blood pressure. The Groningen ART-cohort study and the PGS study used identical follow-up protocols, assessments and questionnaires. Also, routines in both studies were identical. Parents were posted regularly with a newsletter on the study's progress and invited with similar letters to participate in a follow-up round. The routines also implied that the parts of the assessment were always carried out in the same order: cognitive tests, blood pressure measurement, neurological examination, anthropometric assessment and second blood pressure measurement. Follow-up of both studies ran simultaneously and was conducted by the same team of researchers.

To evaluate cognitive development, the Kaufman Assessment Battery for Children, Second Edition (KABC-II) was used.[26] The KABC-II is an American psychological diagnostic test, which is also highly suitable for the assessment of cognitive development of European children and adolescents aged between 3 and 18 years.[26] A total IQ score was calculated on the basis of scores in four domains: sequential processing, evaluating short-term memory; learning ability associated with long-term memory; simultaneous processing, addressing visual processing of information; and planning ability evaluating the competence of logical reasoning. Raw IQ scores of the domain scores and Total IQ were normalized into global scores. US norms were used, because Dutch norms are lacking.[26] The proportion of children with normalized IQ scores below −1 SD, was calculated, a cut-off score that is considered to be clinically relevant.

Neurological condition of the 4-year-old children was evaluated with the Hempel assessment, an age-specific neurological examination to detect MND at preschool age.[27] It assesses neurological function in five domains: fine motor function, gross motor function, muscle tone and posture, reflexes and visuomotor function.[28] A domain can be scored as typical or deviant. The assessment results in a clinical classification: neurologically normal, simple MND, complex MND or major neurological dysfunction, the latter denoting the presence of a distinct neurological syndrome, such as cerebral palsy. The classification of the two forms of MND is based on the number of deviant domains. Complex MND implies the presence of two or more deviant domains. It is the clinically relevant form of MND and is associated with behavioural and learning disorders. Simple MND means one deviant domain; it reflects a normal, but non-optimal neurological condition. Neurologically, normal implies the absence of neurological dysfunction, i.e. the absence of deviant domains.[29] The assessments were carried out by trained assessors supervised by a neurodevelopmental expert (MHA).

The assessment of anthropometrics consisted of the measurement of height, weight, triceps skinfold and subscapular skinfold. The height (in cm) of the children was measured twice using a stadiometer (Seca, Germany). The proportion of children with a height below −1 SD, taking into account sex and age was calculated.[30] The body weight (kg) of the children was measured twice with an electronic weighing scale (Radwag, Random, Poland). The proportion of children having a weight-for-height below −1 SD of the Dutch reference values, taking into account sex and age, was calculated.[30] Triceps and subscapular skinfold thickness (in cm) were each measured three times, alternating between the two skinfolds, on the non-dominant side of the child, using a Servier Calliper. The mean of the two or three anthropometric measurements was used for further calculations.

Blood pressure (mmHg) was measured using an automated blood pressure monitor (Datascope Accutorr plus, Mahwah, NJ, USA). The appropriate cuff size was used: 'child' (arm circumference 13.8–21.5 cm) or 'small adult' (20.5–28.5 cm). Blood pressure was measured at the non-dominant arm whereas the child was seated, with the arm on the lap. Blood pressure measurement was carried out in duplicate. The mean of the two readings was used for further calculations. This mean was used to calculate blood pressure percentiles based on the standards of the US National High Blood Pressure Education Program.[31] The blood pressure percentiles take sex, height and age into account. High blood pressure is defined as a systolic or diastolic percentile above the 95th percentile.

Statistical analysis

To estimate group differences in fertility parameters and parental characteristics, Fisher's exact test, Mann–Whitney U test and student's t-test were used. Mixed-effects models were used for birth characteristics and child characteristics, to account for within family correlation.

The following numerical outcome response variables of children at the age of 4 years were analysed with a linear mixed-effects model: height, weight, triceps skin thickness, subscapular skin thickness, systolic blood pressure, diastolic blood pressure, sequential IQ, learning IQ, simultaneous IQ, knowledge IQ and total IQ. The mother was incorporated as a random subject effect to model the possible correlation between the outcome variables for children from the same family. Singleton and twin status was included as a fixed effect in the mixed effect model. The first set of analyses were unadjusted, the second set of analyses were corrected for IVF/ICSI, TTP, gestational age, small for gestational age (SGA), low Apgar score at 5 min (<7), education level of the parents and body mass index (BMI) of the mother, to obtain an estimate of the unbiased, direct (i.e. unmediated) effect of twin-status outcome. We considered gestational age, SGA, and low Apgar score as potential mediators, TTP, educational level of the parents and maternal body mass index (BMI) as potential confounders. The adjusted analyses for blood pressure also included correction for the potential confounders height, weight and sex, in line with literature.[31]

For the binary response outcome variables: IQ scores below −1 SD, weight-for-height below −1 SD, height below −1 SD, high blood pressure, normal MND, simple MND and complex MND a generalized linear mixed-effects model with the logit link function was used. The estimation was carried out with generalized estimating equation (or by logistic multiple linear regression model when the multilevel model did not prove to be valid) and the cluster variable was determined by the mother again. The robust estimator was used together with an exchangeability working correlation matrix. The effect of singleton and twin status in the first set of analyses was unadjusted for other variables. The second set of analyses were corrected for IVF/ICSI, TTP, gestational age, SGA, low Apgar score at 5 min (<7), education level of the parents and BMI of the mother. The type 3 generalized score statistics were applied to determine the P-value for the twin effect. An additional analysis, in which gestational age was not included in the adjustment was, was conducted afterwards (Supplementary Table S1).

Results of the mixed-effect analyses are expressed as marginal estimated mean values and mean difference values with their 95% confidence intervals (mean values, 95% CI). A non-central t-test power analysis was conducted to estimate the minimal detectable effect size for singletons and twins separately or together based on total IQ. The current sample sizes allowed the detection of a clinically relevant difference of 7.5 points in total IQ between twins and singletons with 80% power ($\alpha = 0.05$). Our sample sizes allowed the detection of a one-half SD difference in the other IQ scores, which we considered clinically relevant. In addition, we conducted a two-sample t-test power analysis for blood pressure. Assuming a SD of 8 mmHg, 42 children had to be included in each group to detect a difference of 5 mmHg and reach a power of 80%.[24,25]

SPSS software, version 22 (IBM Corp., USA) was used for all statistical analyses. $P < 0.05$ was considered to be statistically significant.

Results

Parental and infant characteristics

The ART cohort and PGS study had a similar proportion of twins (29% versus 26%; Fisher's exact test; non-significant). Background characteristics of twins and singletons are presented in Table 1. Parental characteristics, fertility parameters and child characteristics were similar, but birth characteristics differed between the groups. Twins had a lower birthweight ($P < 0.001$) and shorter gestational age at birth. Furthermore, twins were more often delivered by Caesarean section, born preterm and SGA compared with singletons. Note that within the twin group only three twin pairs were monozygotic.

Table I: Prenatal, perinatal, neonatal and demographic characteristics of singletons and twins.

Characteristics	Singletons n=103	Twins n=48	p-value
Parental characteristics			
Maternal age at conception (years), mean (σ)	34.6 (3.9)	34.5 (3.2)	NS
High education level mother, n (%)[b]	43 (42)	26 (54)	NS
Paternal age at conception (years), mean (σ)[a]	37.8 (5.9)	37.7 (3.8)	NS
High education level father, n (%)[a,b]	46 (46)	24 (52)	NS
Maternal BMI at conception, median (range)[a]	23.5 (15.5-42.5)	22.6 (19.0-29.3)	NS
Smoking mother during pregnancy, n (%)[a]	8 (8)	0 (0)	NS
Gestational Diabetes, n (%)[a]	2 (2)	0 (0)	NS
Pregnancy induced hypertension, n (%)[a]	17 (17)	10 (21)	NS
Fertility parameters			
ICSI, n (%)	58 (56)	26 (54)	NS
TTP in years, median (range)	4.1 (0.1-13.3)	4.5 (0.2-10.7)	NS
Birth characteristics			
Gestational age (weeks), mean (σ)	39.1 (1.8)	36.3 (1.9)	(NaN)
Preterm birth (<37 weeks), n (%)	26 (25)	28 (58)	**<.007**
Birthweight (grams), mean (σ)	3378 (567)	2480 (441)	**<.001**
Small for gestational age, n (%)[c]	0 (0)	4 (8)	**.036**
Caesarean section, n (%)	29 (28)	27 (56)	**.017**
Apgar score 5 min <7, n (%)[a]	19 (19)	14 (30)	NS
Child characteristics			
Male sex, n (%)	53 (51)	31 (65)	NS
Firstborn, n (%)	69 (67)	30 (63)	NS
DZ twin, n (%)	n.a.	42 (88)	n.a.
Age (months) at examination, median (range)	49.5 (47.5-70.2)	49.4 (48.1-60.8)	NS

Mann-Whitney U-tests, student t-tests and Fisher's exact tests were used to estimate group differences for parental characteristics and fertility parameters. Mixed effects models were used for birth characteristics and child characteristics, to take into account within-family correlation. Statistically significant numbers ($p < 0.05$) are displayed in bold. Values are number (percentage) or median (range). BMI = Body Mass Index; DZ = Dizygotic; ICSI = Intracytoplasmic Sperm Injection; n.a. = not available; NaN = not a number; NS = non-significant; TTP = Time To Pregnancy.

[a] Missing data in the two groups: Apgar score 5 min, n=2; Gestational diabetes, n=1; High education level father, n=5; Maternal BMI at infant conception, n=9; Paternal age at infant conception, n=21; Pregnancy induced hypertension, n=1; Smoking mother during pregnancy, n=1.

[b] Higher vocational education or University education.

[c] Birthweight for gestational age is < -2 standard deviations compared with the Dutch reference population (Dutch reference tables, perinatal Registration Netherlands).

Neurodevelopmental outcome

Total IQ scores of twins were lower than of singletons (mean values [95% CI]: twins 103.5 [99.8 to 107.3] and singletons 108.9 [106.7 to 111.1]), with a significant mean difference (−5.4 [−9.7 to −1.0]). In addition, twins had more often a total IQ score below −1 SD compared with singletons (27% versus 11%) (Table 2). Simultaneous IQ of twins was lower than that of singletons (twins 109.1 [105.1 to 113.1]; singletons 114.7 [112.3 to 117.2]), whereas sequential, learning and knowledge IQ scores of twins were similar to those of singletons. After adjustment, the differences in IQ scores between twins and singletons disappeared, except for twins having more often a sequential IQ score below −1 SD than singletons (Table 3). Neurological condition of twins did not differ from that of singletons, also after adjustment (Tables 2 and 3).

Anthropometrics

Weight (mean values, 95% CI: twins 16.9 [16.0 to 17.9]; singletons 18.6 [18.1 to 19.1]) and height (twins 105.9 [104.0 to 107.7]; singletons 108.8 [107.9 to 109.8]) were significantly lower in twins than in singletons. Height below −1 SD and weight for height below −1 SD, however, did not differ between twins and singletons. Moreover, triceps and subscapular skinfold thickness in the two groups were similar (Table 2). After adjustment, all anthropometric values, including weight and height, were similar between twins and singletons (Table 3).

Blood pressure

Unadjusted mixed-model analyses indicated that blood pressure of twins was similar to that of singletons (Table 2) also after adjustment for confounders (Table 3). Also, high blood pressure rates were similar in the two groups (Table 2).

Table II: Cognitive and neurological outcome, anthropometrics and blood pressure in singletons and twins: results of the unadjusted mixed-effect model analyses

Outcomes	Singletons (n=103)	Twins (n=48)	Mean difference [CI]	p-value
Cognitive outcome				
Sequential IQ, mean [CI]	99.2 [97.0; 101.5]	95.0 [91.0; 98.9]	-4.2 [-8.8; 0.3]	NS
Sequential IQ score below -1SD, n (%)[b]	27 (26)	21 (44)		NS
Learning IQ, mean [CI]	98.5 [96.4; 100.6]	96.6 [93.5; 99.7]	-1.9 [-5.8; 1.9]	NS
Learning IQ score below -1SD, n (%)[b]	21 (20)	14 (29)		NS
Simultaneous IQ, mean [CI]	114.7 [112.3; 117.2]	109.1 [105.1; 113.1]	-5.6 [-10.4; -1.0]	**0.019**
Simultaneous IQ score below -1SD, n (%)[b]	9 (9)	11 (23)		NS
Knowledge IQ, mean [CI]	111.1 [108.5; 113.7]	106.4 [101.6; 111.3]	-4.7 [-10.2; 0.8]	NS
Knowledge IQ score below -1SD, n (%)[b]	8 (8)	5 (10)		NS
Total IQ, mean [CI]	108.9 [106.7; 111.1]	103.5 [99.8; 107.3]	-5.4 [-9.7; -1.0]	**0.016**
Total IQ score below -1SD, n (%)[b]	11 (11)	13 (27)		**0.047**
Neurological outcome				
Normal, n (%)	60 (58)	23 (48)		NS
Simple MND, n (%)	14 (14)	10 (21)		NS
Complex MND, n (%)	29 (28)	15 (31)		NS
Major neurological dysfunction, n (%)	0 (0)	0 (0)		N.A.
Anthropometrics				
Weight (kg), mean [CI][a]	18.6 [18.1; 19.1]	16.9 [16.0; 17.9]	-1.7 [-2.7; -0.6]	**0.003**
Weight for height below -1SD, n (%)[a,c]	6 (7)	7 (16)		NS
Standing height (cm), mean [CI][a]	108.8 [107.9; 109.8]	105.9 [104.0; 107.7]	-2.9 [-5.0; -0.8]	**0.007**
Height below -1SD, n (%)[a,c]	11 (11)	11 (24)		NS
Triceps skinfold (cm), mean [CI][a]	1.08 [1.01; 1.16]	1.01 [0.875; 1.15]	-0.07 [-0.23; 0.08]	NS
Subscapular skinfold (cm), mean [CI][a]	0.558 [0.509; 0.608]	0.540 [0.451; 0.629]	-0.018 [-0.119; 0.084]	NS
Blood pressure				
SBP, mean [CI] mmHg[a]	101.8 [100.2; 103.4]	102.3 [99.9 ; 104.8]	0.5 [-2.4; 3.5]	NS
DBP, mean [CI] mmHg[a]	63.5 [62.0; 65.0]	63.4 [60.8 ; 65.9]	-0.1 [-3.1; 2.8]	NS
High Blood Pressure, n (%)[a,d]	16 (17)	10 (24)		NS

Statistically significant differences ($p<0.05$) are displayed in bold. BP = Blood Pressure; DBP; Diastolic Blood Pressure; IQ = Intelligence Quotient; MND = Minor Neurological Dysfunction; NS = non-significant; SBP = Systolic Blood Pressure.

[a] Missing data in the two groups: standing height, n=5; height under 1SD, n=5, DBP n=11; High BP, n=12; SBP, n=11; Subscapular skinfold, n=14; Triceps skinfold, n=9; weight, n=7; weight for height under 1SD, n=7.

[b] Scoring below -1SD on age based norm values of the Kaufmann test (Kaufman, 2004).

[c] Scoring below -1SD on age based Dutch norm values (Schönbeck and van Buuren, 2010).

[d] High BP is defined as a SBP- or DBP percentile above the 95th percentile according to the standards of the U.S. National High BP Education Program (Falkner and Daniels, 2004). The BP percentiles take sex, height and age into account.

Table III: Health and development of twins and singletons: results of the adjusted mixed-effect model analyses

Response variable	Singletons (n=96)[a]	Twins (n=46)[a]	Mean difference [CI]	P-value
Cognitive outcomes				
Sequential IQ, mean [CI]	99.0 [91.8; 106.1]	93.5 [86.5; 100.5]	-5.5 [-11.2; 0.2]	NS
Sequential IQ score below -1SD, n (%)[c]	27 (28)	20 (46)		0.030[h]
Learning IQ, mean [CI]	90.4 [83.1; 97.7]	89.9 [82.8; 96.9]	-0.5 [-6.2; 5.1]	NS
Learning IQ score below -1SD, n (%)[c]	20 (21)	14 (30)		NS[g]
Simultaneous IQ, mean [CI]	100.1 [92.4; 107.8]	97.4 [89.8; 105.1]	-2.7 [-9.1; 3.8]	NS
Simultaneous IQ score below -1SD, n (%)[c]	9 (9)	10 (22)		NS[g]
Knowledge IQ, mean [CI]	106.1 [99.5; 112.7]	101.2 [93.7; 108.8]	-4.8 [-12.2; 2.5]	NS
Knowledge IQ score below -1SD, n (%)[c]	8 (8)	5 (11)		NS[h]
Total IQ, mean [CI]	98.2 [91.9; 104.6]	94.2 [87.4; 100.9]	-4.0 [-10.2; 2.0]	NS
Total IQ score below -1SD, n (%)[c]	11 (11)	12 (26)		NS[h]
Anthropometrics				
Weight, mean [CI][b]	18.8 [17.2; 20.4]	17.8 [16.2; 19.4]	-1.0 [-2.4; 0.5]	NS
Weight for height below -1SD, n (%)[b,d]	6 (6)	6 (14)		NS[g]
Standing height, mean [CI][b]	108.6 [106.0; 111.2]	107.4 [104.6; 110.1]	-1.2 [-3.7; 1.3]	NS
Height below -1SD, n (%)[b,d]	11 (12)	10 (23)		NS[h]
Triceps skin thickness, mean [CI][b]	1.26 [1.06; 1.45]	1.21 [1.00; 1.43]	-0.04 [-0.2; 0.2]	NS
Subscapular skin thickness, mean [CI][b]	0.703 [0.549; 0.856]	0.687 [0.533; 0.840]	-0.016 [-0.2; 0.1]	NS
Blood pressure				
Systolic blood pressure, mean [CI][b]	101.8 [96.3; 107.3]	105.1 [99.9; 110.4]	3.3 [-0.9; 7.4]	NS
Diastolic blood pressure, mean [CI][b]	62.6 [57.3; 67.9]	64.9 [59.8; 70.0]	2.3 [-1,7; 6.3]	NS
High Blood Pressure, n (%)[b,e]	15 (17)	10 (25)		NS[g]
Neurological outcome				
Normal MND, n (%)	54 (56)	22 (48)		NS[f]
Simple MND, n (%)	14 (15)	10 (22)		NS
Complex MND, n (%)	28 (29)	14 (30)		NS

Statistically significant differences ($p < 0.05$) are displayed in bold.
IQ = Intelligence Quotient; MND = minor neurological dysfunction; NS = non-significant.
In the adjusted analyses we corrected for IVF/ICSI, time to pregnancy, gestational age, SGA, low Apgar score at 5 min (<7), education level parents and Body Mass Index of the mother. The adjusted analyses for blood pressure also included correction for height, weight and sex
[a] Nine children, 7 singletons and 1 twin pair, were not included because data on maternal BMI was missing.
[b] Missing data in the two groups: standing height, n=5; height under 1SD, n=5, DBP n=11; High BP, n=12; SBP, n=11; Subscapular skinfold, n=14; Triceps skinfold, n=9; weight, n=6; weight for height under 1SD, n=6.
[c] Scoring below -1SD on age based norm values of the Kaufmann test (Kaufman, 2004).
[d] Scoring below -1SD on age based Dutch norm values (Schönbeck and van Buuren, 2010).
[e] High BP is defined as a SBP- or DBP percentile above the 95th percentile according to the standards of the U.S. National High BP Education Program. The BP percentiles take sex, height and age into account (Falkner and Daniels, 2004).
[f] Resulting from the generalized linear mixed model with all selected background variables except small for gestational age (SGA).
[g] Logistic multiple linear regression model due to convergence issues with the generalized linear mixed model.
[h] All confidence intervals for the estimated marginal means were maximal width due to the small numbers in the cells. Results should therefore be interpreted with caution.

Discussion

The present study demonstrated that 4-year-old IVF twins have a significantly lower total IQ, a lower body weight and a smaller height than 4-year-old IVF singletons. After adjustment for confounders, and taking into account mediators such as gestational age and SGA, development and health of 4-year-old IVF twins did not differ from that of 4-year-old IVF singletons: cognitive development, neurological outcome, anthropometrics and blood pressure of 4-year-old IVF twins were similar to that of IVF singletons.[16-19]

Our findings that development and health of IVF twins is less favourable than that of singletons is in line with the large body of literature on twins and singletons in general.[4,5,7-9,32] The difference in outcome may largely be attributed to differences in mediating perinatal complications, a difference that was also present in our groups.

We found that the total IQ scores of the IVF twins were about one third of a SD lower than those of the IVF singletons, a difference that corresponds to the IQ differences found in general between twins and singletons.[33] The overall lower IQ scores resulted in a doubling among twins of the proportion of children with a total IQ score below −1 SD. This finding is clinically highly relevant as a lower IQ at preschool age predicts to a large extent a lower IQ at the second half of school age.[14] Simultaneous IQ was also lower in IVF twins. We did not, however, see a difference in simultaneous IQ below −1 SD between IVF twins and IVF singletons, as some twins scored relatively low whereas others scored relatively high. Our findings of a difference in cognitive outcome between IVF twins and IVF singletons also matches with the findings of Bonduelle et al. who found that 2-year-old IVF and ICSI twins scored significantly lower on the mental developmental index of the Bayley Scale of Infant development than IVF and ICSI singletons.[9] In the Bonduelle study, the comparison of twins and singletons was not adjusted for effect mediators and confounders. Our study, however, suggests that the differences in cognitive development disappear after adjustment. This indicates that the less favourable cognitive outcomes may be mainly attributed to the higher rate of perinatal adversities in twins than to an independent twin effect.

Pinborg et al. reported that IVF twins more often had speech problems than IVF singletons; however, in that study, only birth weight was taken into account as effect mediator.[10] In line with another study by Pinborg et al., no differences were found in neurological outcome between IVF-twins and IVF-singletons.[11] The present study, however, was too small to detect differences in the prevalence of cerebral palsy.

We found that twins born after IVF had a significantly smaller height and lower weight than singletons born after IVF, resulting in a statistically non-significant doubling of children with a height and weight-for-height below 1 SD. It has been well documented that a smaller height and weight at preschool age is associated with smaller body proportions in later life.[14,34] The present study showed, however, that the increased risk for smaller anthropometrics disappeared when the modifying effects of gestational age and SGA were taken into account. Saunders et al. described a similar finding in 2-year-old twins in general: the mean weight of twins was not different from that of 2-year-old singletons after correction for prematurity.[35]

Our study also indicated that IVF twins and IVF singletons have similar blood pressure, a finding that corresponds to that of De Geus et al. who reported that, in general, blood pressures of adult twins matches that of adult singletons.[15]

Our results indicate that IVF twins compared with IVF singletons are at increased risk for a less favourable cognitive outcome, have a lower body weight and a smaller standing height. The effect is, as in general populations of twins, largely mediated by the increased risk of twins for adverse perinatal outcome.[32] This implies that IVF per se does not alter the increased risk for impaired development and health associated with twin-ship.

Strengths and weaknesses

To the best of our knowledge, this is the first study addressing in detail neurodevelopmental and cardiometabolic condition at 4 years of IVF twins compared with that of IVF singletons. Other strengths are the recruitment of couples during pregnancy preventing selection bias, the relatively small attrition (less than 15% after 4 years of follow-up) and the mixed-effects models where the mother was incorporated as a random subject effect to model the possible correlation between the outcome variables for children from the same family. The covariates used in the adjusted mixed effect models may be considered another strength of the study, covariates that included the potential mediator TTP and the potential confounder maternal BMI. We selected the covariates on an a-priori bases, in keeping with the literature, our prior knowledge of both studies and the sample sizes. We limited the number of covariates in the adjusted analyses to avoid overstratification. It may be regarded a strength but also a limitation of the study that we pooled two studies. The benefit is the increase in statistical power. Yet, the few background differences between the two groups may be a pitfall as it may have introduced some bias. Nevertheless, we consider the latter risk relatively small, as the background differences were limited and follow-up assessments, follow-up period and the team of researchers were similar.

A limitation of the study is that we did not distinguish between monozygotic and dizygotic twins. Monozygotic twins may share one placenta or each twin partner may have its own, whereas dizygotic twins always have their own placenta. Monozygotic twins who share one placenta are at higher risk for adverse obstetric outcomes than twins who have their own placenta.[36] Most of the assisted reproduction technique twins in general and also in our study are dizygotic owing to dual embryo transfer.

Another limitation of the study is that all blood pressure measurements were carried out on one day. A 24-h telemetric blood pressure measurement or multiple blood pressure measurements on several days would be more accurate. Because of logistic reasons, time and costs, multiple measurements on multiple days were not possible.

In conclusion, the results of the present study suggest that neurological condition and blood pressure of 4-year-old IVF twins is similar to those of their IVF singleton peers. In addition, our study indicates that 4-year-old IVF twins have a significantly lower total IQ, lower body weight and lower height than 4-year old IVF singletons. The less favourable outcome of the twins could be attributed to mediators and confounders. After adjusting for these factors, development and health of 4-year-old IVF twins was similar to that of 4-year-old IVF singletons. This suggests that the generally observed increased risk for less favourable health and development in twins also holds true for IVF twins, and is not altered by IVF per se.

To confirm that IVF does not alter the increased risk of twins for less optimal development and health future research is needed in which studies compare the long-term outcome of naturally conceived twins with those conceived with assisted reproduction techniques, as our findings cannot be considered conclusive and should be interpreted with caution. The scientific and clinical community would especially be helped by studies that include the evaluation of individual, longitudinal trajectories of growth and development. Last, but not least, our findings emphasize the importance of elective single embryo transfer to reduce the worldwide increasing twin birth rate resulting from assisted reproduction techniques, as health and development of twins is less favourable compared with that of singletons, most likely because of the increased risk of twins for perinatal complications.

Acknowledgements

We are greatly thankful to the parents and children who participated in the study, and Linze Dijkstra and Loes de Weerd for their technical assistance.

References

1. Collins J. Global epidemiology of multiple birth. *Reprod Biomed Online* 2007;**15**:45-52.
2. Calhaz-Jorge C, de Geyter C, Kupka MS, de Mouzon J, Erb K, Mocanu E, Motrenko T, Scaravelli G, Wyns C, Goossens V. Assisted reproductive technology in Europe, 2012: results generated from European registers by ESHRE. *Hum Reprod* 2016;**31**:1638-52.
3. Spellacy WN, Handler A, Ferre CD. A case-control study of 1253 twin pregnancies from a 1982-1987 perinatal data base. *Obstet Gynecol* 1990;**75**:168-71.
4. Adler-Levy Y, Lunenfeld E, Levy A. Obstetric outcome of twin pregnancies conceived by in vitro fertilization and ovulation induction compared with those conceived spontaneously. *Eur J Obstet Gynecol Reprod Biol* 2007;**133**:173-8.
5. Baxi A, Kaushal M. Outcome of twin pregnancies conceived after assisted reproductive techniques. *J Hum Reprod Sci* 2008;**1**:25-8.
6. Bensdorp AJ, Hukkelhoven CW, van der Veen F, Mol BW, Lambalk CB, van Wely M. Dizygotic twin pregnancies after medically assisted reproduction and after natural conception: maternal and perinatal outcomes. *Fertil Steril* 2016;**106**:371-7.
7. Botting BJ, Davies IM, Macfarlane AJ. Recent trends in the incidence of multiple births and associated mortality. *Arch Dis Child* 1987;**62**:941-50.
8. Pinborg A. IVF/ICSI twin pregnancies: risks and prevention. *Hum Reprod Update* 2005;**11**:575-93.
9. Bonduelle M, Ponjaert I, Steirteghem AV, Derde MP, Devroey P, Liebaers I. Developmental outcome at 2 years of age for children born after ICSI compared with children born after IVF. *Hum Reprod* 2003;**18**:342-50.
10. Pinborg A, Loft A, Schmidt L, Andersen AN. Morbidity in a Danish national cohort of 472 IVF/ICSI twins, 1132 non-IVF/ICSI twins and 634 IVF/ICSI singletons: health-related and social implications for the children and their families. *Hum Reprod* 2003;**18**:1234-43.
11. Pinborg A, Loft A, Schmidt L, Greisen G, Rasmussen S, Andersen AN. Neurological sequelae in twins born after assisted conception: controlled national cohort study. *BMJ* 2004;**329**:311.
12. Buckler JM, Green M. A comparison of the early growth of twins and singletons. *Ann Hum Biol* 2004;**31**:311-32.
13. Chaudhari S, Bhalerao MR, Vaidya U, Pandit A, Nene U. Growth and development of twins compared with singletons at ages one and four years. *Indian Pediatr* 1997;**34**:1081-6.
14. Silva PA, Crosado B. The growth and development of twins compared with singletons at ages 9 and 11. *Aust Paediatr J* 1985;**21**:265-7.
15. De Geus EJ, Posthuma D, Ijzerman RG, Boomsma DI. Comparing blood pressure of twins and their singleton siblings: being a twin does not affect adult blood pressure. *Twin Res* 2001;**4**:385-91.
16. Hutton JL, Pharoah PO, Cooke RW, Stevenson RC. Differential effects of preterm birth and small gestational age on cognitive and motor development. *Arch Dis Child Fetal Neonatal Ed* 1997;**76**:F75-81.
17. Lagerström M, Bremme K, Eneroth P, Magnusson D. School performance and IQ-test scores at age 13 as related to birth weight and gestational age. *Scand J Psychol* 1991;**32**:316-24.
18. Petrini JR, Dias T, McCormick MC, Massolo ML, Green NS, Escobar GJ. Increased risk of adverse neurological development for late preterm infants. *J Pediatr* 2009;**154**:169-76.
19. Wells JC, Chomtho S, Fewtrell MS. Programming of body composition by early growth and nutrition. *Proc Nutr Soc* 2007;**66**:423-34.
20. Middelburg KJ, Heineman MJ, Bos AF, Pereboom M, Fidler V, Hadders-Algra M. The Groningen ART cohort study: ovarian hyperstimulation and the in vitro procedure do not affect neurological outcome in infancy. *Hum Reprod* 2009;**24**:3119-26.
21. Mastenbroek S, Twisk M, van Echten-Arends J, Sikkema-Raddatz B, Korevaar JC, Verhoeve HR, Vogel NE, Arts EG, de Vries JW, Bossuyt PM, Buys CH, Heineman MJ, Repping S, van der Veen F. In vitro fertilization with preimplantation genetic screening. *N Engl J Med* 2007;**357**:9-17.
22. Schendelaar P, Middelburg KJ, Bos AF, Heineman MJ, Kok JH, La Bastide-Van Gemert S, Seggers J, van den Heuvel ER, Hadders-Algra. The effect of preimplantation genetic screening on neurological, cognitive and behavioural development in 4-year-old children: follow-up of a RCT. *Hum Reprod* 2013;**28**:1508-18.
23. Schendelaar P, Van den Heuvel ER, Heineman MJ, La Bastide-Van Gemert S, Middelburg KJ, Seggers J, Hadders-Algra M. Increased time to pregnancy is associated with less optimal neurological condition in 4-year-old singletons, in vitro fertilization itself is not. *Hum Reprod* 2014;**29**:2773-86.
24. Seggers J, Haadsma ML, Bastide-van Gemert S, Heineman MJ, Kok JH, Middelburg KJ, Roseboom TJ, Schendelaar P, van den Heuvel ER, Hadders-Algra M. Blood pressure and anthropometrics of 4-y-old children born after preimplantation genetic screening: follow-up of a unique, moderately sized, randomized controlled trial. *Pediatr Res* 2013;**74**:606-14.
25. Seggers J, Haadsma ML, La Bastide-Van Gemert S, Heineman MJ, Middelburg KJ, Roseboom TJ, Schendelaar P, Van den Heuvel ER, Hadders-Algra M. Is ovarian hyperstimulation associated with higher blood pressure in 4-year-old IVF offspring? Part I: multivariable regression analysis. *Hum Reprod* 2014;**29**:502-9.
26. Kaufman AS, Kaufman NL. *Kaufman Assessment Battery for Children*, 2nd edn, 2004. AGS Publishing, Circle Pines, The United States of America.
27. Hempel MS. *The neurological examination technique for toddler-age*, 1st edn, 1993. University of Groningen, Groningen, The Netherlands.
28. Hadders-Algra M. The neuromotor examination of the preschool child and its prognostic significance. *Ment Retard Dev Disabil Res Rev* 2005;**11**:180-8.
29. Hadders-Algra M. Two distinct forms of minor neurological dysfunction: perspectives emerging from a review of data of the Groningen Perinatal Project. *Dev Med Child Neurol* 2002;**44**:561-71.
30. Schönbeck Y, van Buuren S. *Vijfde Landelijke Groeistudie*, 2010. TNO, Leiden, The Netherlands.
31. Falkner B, Daniels SR. Summary of the fourth report on the diagnosis, evaluation, and treatment of high blood pressure in children and adolescents. *Hypertension* 2004;**44**:387-8.
32. Sutcliffe AG, Derom C. Follow-up of twins: health, behaviour, speech, language outcomes and implications for parents. *Early Hum Dev* 2006;**82**:379-86.
33. Voracek M, Haubner T. Twin-singleton differences in intelligence: a meta-analysis. *Psychol Rep* 2008;**102**:951-62.
34. Estourgie-van Burk GF, Bartels M, Van Beijsterveldt TC, Delemarre-van de Waal HA, Boomsma DI. Body size in five-year-old twins: heritability and comparison to singleton standards. *Twin Res* 2006;**9:**646-5.
35. Saunders K, Spensley J, Munro J, Halasz G. Growth and physical outcome of children conceived by in vitro fertilization. *Pediatrics* 1996:**97**;688-92.
36. Minakami H, Sayama M, Honma Y, Matsubara S, Koike T, Sato I, Uchida A, Eguchi Y, Momoi M, Araki S. Lower risks of adverse outcome in twins conceived by artificial reproductive techniques compared with spontaneously conceived twins. *Hum Reprod* 1998;**13**:2005-8.